Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years

Anthony B. Miller, MB, FRCP; Cornelia J. Baines, MD, MSc; Teresa To, PhD; Claus Wall, MSc

Objectives: To evaluate the efficacy of the combination of annual screening with mammography, physical examination of the breasts and the teaching of breast self-examination in reducing the rate of death from breast cancer among women aged 40 to 49 years on entry.
Design: Individually randomized controlled trial.
Setting: Fifteen urban centres in Canada with expertise in the diagnosis and treatment of breast cancer.
Participants: Women with no history of breast cancer and no mammography in the previous 12 months were randomly assigned to undergo either annual mammography and physical examination (MP group) or usual care after an initial physical examination (UC group). The 50 430 women enrolled from January 1980 through March 1985 were followed for a mean of 8.5 years.
Data collection: Derived from the participants by initial and annual self-administered questionnaires, from the screening examinations, from the patients' physicians, from the provincial cancer registries and by record linkage to the Canadian National Mortality Data Base. Expert panels evaluated histologic and death data.
Main outcome measures: Rates of referral from screening, rates of detection of breast cancer from screening and from community care, nodal status, tumour size, and rates of death from all causes and from breast cancer.
Results: Over 90% of the women in each group attended the screening sessions or returned the annual questionnaires, or both, over years 2 to 5. The characteristics of the

From the Department of Preventive Medicine and Biostatistics, University of Toronto, Toronto, Ont.

Canadian National Breast Screening Study investigators: ***centre directors,*** *Drs. Alan A. Bassett, D.C. Gordon Bethune, David M. Bowman, Haydn Bush, Jacques Cantin, Luc Dêschenes, James E. Devitt, D. Neil Graham, Gregory Hislop, Alan W. Lees, Bernard M. Leféhvre, Leo Mahoney, S. Edward O'Brien, Antoine Simard and Wally J. Temple;* ***surgeons,*** *Drs. C. Paul Armstrong, Robert M. Baird, Wayne Beecroft (deceased), William J. Buie, Robert Bury, Chris D.J. Chadwick, W. Gordon Chipperfield, Donald Currie, Gary J. Dewar, Maurice Falardeau, Gerald J. Francis, Morris H. Friedman, Nenad Gagic, David Girvin, Donald J. Hamilton, H. Reginald Harse, Irving Koven, Urve Kuusk, Alan B. McCarten, John McCredie, R. Douglas Marriott, William O. Onerheim, André Péloquin, Claude Potvin, Robert E. Pow, James Purves, Patricia M. Rebbeck, Jean Robert, André Robidoux, J. Trevor Sandy, Saul Sidlofsky, Elin R. Sigurdson, Bernard J. Steele, Robert M. Stone, J. Bernard Taillefer, T. Kenneth Thorlakson and Gordon K. Thorson;* ***radiologists,*** *Drs. Luc Audet, Bruce L. Bird, Margaret J. Burns, Bernice Capusten, William R. Castor, Gabriel M. Cooke, Catherine M. Copeland, James W. Davidson, Gerald E.D. Davis, J.E. Leo Desautels, René L. Desmarais, Laszlo A. Fried, André Grégoire, George Hardy, Patricia Hassell, Guy Hébert, Roberta Jong, Sheila M. Kelly, Jacques Ladouceur (deceased), Jean Laperrière, J. Donald Longley, Roberta N. Ludwig, John H.M. MacGregor, John McCallum, J. Stewart Manchester, Terry Minuk, Hugh F. Morrish, Helmut A. Mueller, Denise Ouimet-Oliva, Norman L. Patt, Michel Petitclerc, Peter Poon, Odile Prosmanne, John W. Radomsky, Pasteur Rasuli, Jean-L. Robillard, Bernard J. Shapiro, Saul L. Share, Imre S. Simor, R. Keith Sparrow, Harold K. Standing, William J. Weiser and Arthur H. Zalev;* ***pathologists,*** *Drs. Fred Alexander, Yvan Boivin, Noel Cooter, John Danyluk, David Dawson, Terrence J. D'Souza, Maha Jabi, Simon Jacob, Janice Safneck, Walter Schurch, Harry Strawbridge, D. Ian Turnbull, René Vauclair, Ann Worth, Hossein Yazdi and Ismail Zayid.*

Reprint requests to: Department of Preventive Medicine and Biostatistics, University of Toronto, Toronto, ON M5S 1A8

women in the two groups were similar. Compared with the Canadian population, the participants were more likely to be married, have fewer children, have more education, be in a professional occupation, smoke less and have been born in North America. The rate of screen-detected breast cancer on first examination was 3.89 per 1000 in the MP group and 2.46 per 1000 in the UC group; more node-positive tumours were found in the MP group than in the UC group. During years 2 through 5 the ratios of observed to expected cases of invasive breast cancer were 1.26 in the MP group and 1.02 in the UC group. Of the women with invasive breast cancer through to 7 years, 191 and 157 women in the MP and UC groups respectively had no node involvement, 55 and 43 had one to three nodes involved, 47 and 23 had four or more nodes involved, and 38 and 49 had an unknown nodal status. There were 38 deaths from breast cancer in the MP group and 28 in the UC group. The ratio of the proportions of death from breast cancer in the MP group compared with those in the UC group was 1.36 (95% confidence interval 0.84 to 2.21). The survival rates were similar in the two groups. The highest survival rate occurred among women whose cancer had been detected by mammography alone.

Conclusion: The study was internally valid, and there was no evidence of randomization bias. Screening with yearly mammography and physical examination of the breasts detected considerably more node-negative, small tumours than usual care, but it had no impact on the rate of death from breast cancer up to 7 years' follow-up from entry.

Objectifs : Évaluer l'efficacité de la combinaison de tests annuels de dépistage avec mammographie, examen physique des seins et enseignement de l'auto-examen des seins pour réduire le taux de décès dû au cancer du sein chez les femmes de 40 à 49 ans à l'entrée.

Conception : Étude aléatoire contrôlée individuelle.

Cadre : Quinze centres urbains au Canada possédant une compétence en diagnostic et traitement du cancer du sein.

Participantes : Nous avons choisi au hasard des femmes sans antécédents de cancer du sein et n'ayant pas subi de mammographie dans les 12 mois précédents afin qu'elles passent soit une mammographie annuelle avec examen physique (groupe MP) ou reçoivent les soins habituels après l'examen physique initial (groupe UC). Les 50 430 femmes inscrites de janvier 1980 à mars 1985 ont été suivies 8,5 ans en moyenne.

Collection des données : Données obtenues des participantes grâce au questionnaire d'entrée, aux questionnaires annuels autoadministrés et aux examens de dépistage et par l'entremise des médecins traitants, des registres provinciaux de cancérologie et en établissant un lien entre le dossier et la base de données nationale sur la mortalité au Canada. Les données histologiques et de décès ont été évaluées par des groupes de spécialistes.

Principales mesures des résultats : Taux de consultation découlant du dépistage, taux de détection du cancer du sein par dépistage et par les soins de santé communautaire, état ganglionnaire, taille de la tumeur et taux de décès de toutes causes et par cancer du sein.

Résultats : Plus de 90 % des femmes de chaque groupe ont assisté aux séances de dépistage ou ont retourné les questionnaires annuels, ou les deux, de la 2e à 5e année. Dans les deux groupes, les femmes présentaient des caractéristiques analogues. Comparativement à la population canadienne, les participantes étaient plus susceptibles d'être mariées, d'avoir moins d'enfants, d'être plus scolarisées, d'occuper une situation professionnelle, de moins fumer et d'être nées en Amérique du Nord. Le taux de cancers du sein décelés par dépistage au premier examen était de 3,89 par 1 000 dans le groupe MP et de 2,46 par 1 000 dans le groupe UC; nous avons relevé un plus grand nombre de tumeurs à atteinte ganglionnaire dans le groupe MP que dans le groupe UC. De la 2e à la 5e années, les ratios entre les cas observés et prévus de cancers envahissants du sein étaient de 1,26 dans le groupe MP et de 1,02 dans le groupe UC. Parmi les femmes atteintes de cancer envahissant du sein au cours des 7 années, 191 et 157 femmes des groupes MP et UC respectivement ne présentaient aucune atteinte ganglionnaire, 53 et 43 présentaient d'un à trois ganglions atteints, et 47 et 23, au moins quatre ganglions atteints, tandis que pour 38 et 49 l'état ganglionnaire était inconnu. Il y a eu 38 décès par cancer du sein dans le groupe MP et 28 dans le groupe UC. Le ratio des proportions de décès par cancer du sein dans le groupe MP comparativement au taux dans le groupe UC était de 1,36 (intervalle de confiance de 95 %, 0,84 à 2,21). Les taux de survie étaient analogues dans les deux groupes. Le taux de survie le plus élevé a été observé chez les femmes dont le cancer avait été décelé par la mammographie uniquement.

Conclusion : L'étude était valide sur le plan interne et nous n'avons relevé aucune preuve de gauchissement aléatoire. Le dépistage par mammographie annuelle et examen physique des seins a permis de déceler beaucoup plus de petites tumeurs sans atteinte ganglionnaire que les soins habituels, mais n'a eu aucun effet sur le taux de décès par cancer du sein jusqu'à 7 années du suivi à partir de l'entrée.

Part 1 of the Canadian National Breast Screening Study (NBSS) is an individually randomized trial designed to evaluate the efficacy of the combination of annual mammography, physical examination of the breasts and the teaching of breast self-examination in reducing the rate of death from breast cancer among women aged 40 to 49 years on entry.[1]

The NBSS had a prolonged gestation period, which started with a recommendation in 1974 that a study be performed.[2,3] A national planning meeting was held in 1976 and attended by representatives from most of the centres that eventually participated in the study. Pilot studies were performed in 1977. The final design of the study was approved by the National Cancer Institute of Canada and the Department of National Health and Welfare in 1979.

Screening for breast cancer among women aged 40 to 49 years is controversial.[4] When the design of the NBSS was finalized (in 1979) there was no evidence that such screening was effective among women in this age group.[5] Indeed, the Working Group to Review the National Cancer Institute–American Cancer Society Breast Cancer Detection Demonstration Projects had recommended that a trial be conducted to examine this issue.[5] The NBSS was designed to meet this need. Subsequently, the only evidence of the efficacy of screening in reducing the rate of death from breast cancer in this age group has come from the long-term follow-up of the Health Insurance Plan (HIP) study.[6,7] However, the evidence is controversial,[4] and the findings require confirmation in another trial. There is clearly little or no evidence of the effectiveness of mammography in reducing the rate of death from breast cancer, at least in the first 10 years after initiation of screening.[8] The largest trial to date involving women 40 to 49 years of age, the Swedish two-county trial,[9] showed no evidence of a decreased rate of death from breast cancer after 11 years of follow-up.

In this article we concentrate on the findings from the first 7 years of follow-up for women aged 40 to 49 years on entry to the NBSS. Part 2 of the NBSS, reported in the accompanying article (see pages 1477 to 1488 of this issue), evaluated the efficacy of annual mammography over and above annual physical examination of the breasts and the teaching of breast self-examination in reducing the rate of death from breast cancer among women aged 50 to 59 years on entry to the study. Reports on some aspects of the study have appeared, including those on the early phase of recruitment,[10] changes in breast self-examination behaviour,[11] the sensitivity of the screening methods[12-14] and the early results.[15,16]

Methods

Coordination

The study was coordinated at the NBSS Central Office, University of Toronto. During the operation of the study centres the authors were assisted by a full-time national coordinator. Each study centre had a director (usually a physician), a surgeon (sometimes the director of the centre), designated radiologist(s) and a pathologist. The surgeons, radiologists and pathologists attended routine meetings during the screening period.

Centre coordinators were selected from applicants, mostly nurses, with experience in clinic or study management. Each coordinator received intensive training and support from the central office. The study procedures were set out in an operations manual. The national coordinator attended the initial week of each centre's opening, ensuring that study procedures were understood and followed. Thereafter, the national coordinator and two of us (A.B.M. and C.J.B.) regularly visited the centres. Phone contact occurred several times a week between the local coordinators and the national office. Special day-long meetings were held for the coordinators on a regular (at least annual) basis.

Selection of the study centres

The criteria for selecting the study centres were a base population in an area sufficient to recruit the required number of participants and a centre in that area with interest in the study and sufficient professional expertise to comply with the study protocol. The protocol required expertise in mammography and in the diagnosis and treatment of breast cancer. Since breast screening programs were not in operation anywhere in Canada at the time the study was initiated, experience was based on diagnostic mammography. There were no designated training programs in screening available then in North America for radiologists and radiographers, or for physicians or nurse examiners. Instead, special quality-control mechanisms were established for radiation physics and radiology,[17-19] and a protocol was prepared for the local training of the physical examiners.[14,20] Facilities and equipment for modern film-screen mammography were absolute prerequisites. Xerography was not used. The mammography equipment used in the study has been described previously.[21]

Funds were provided in 1980 for 2 screening centres, in 1981 for 3 more and in 1983 for the remaining 10. Screening centres were located in Nova Scotia, Quebec, Ontario, Manitoba, Alberta and British Columbia.

Sample size

The sample size was fixed to determine whether a reduction of 40% in the rate of death from breast cancer would be seen in the intervention group, as compared with the control group. This level of reduction was judged appropriate because it was achieved among women over 50 years of age in the HIP trial at 5 years after entry[6] and it was assumed that with modern mammography a similar reduction in the death rate should occur among women aged 40 to 49 years.[1]

Given the rate of death from breast cancer in Canada among women aged 40 to 49, the total sample would have to be 50 000 women, at an α level of 0.05 and a power of 80% after 5 years of follow-up.[1] In practice, at 5 years the number of deaths from breast cancer was insufficient to achieve the planned power. Therefore, the follow-up was extended for 2 years.

Recruitment of participants

Because the study was an efficacy trial, with the basic planned comparisons internal, we did not attempt to recruit a representative proportion of the eligible female population in any of the study areas.

Participants were recruited by various means. In 5 of the 15 study areas population lists were used to generate personalized letters of invitation. In four of the five (Toronto, Hamilton, Ottawa and London) the lists had been compiled for municipal taxes; about 2 weeks after the letters were sent, staff telephoned those who had so far not requested an appointment. In the fifth area (Halifax) the provincial health insurance office sent letters on our behalf; to protect confidentiality the NBSS staff did not have access to the records, and no reminder or telephone follow-up was possible. In Toronto and Halifax an estimated 25% and 30% respectively of the women approached were enrolled in the study. A mailed invitation was given as the reason for entering the study by 12% of the participants in Toronto, 5% in Hamilton, 13% in Ottawa, 5% in London and 28% in Halifax.[22]

General publicity was used extensively through media contacts, often generated by the local divisions and units of the Canadian Cancer Society. Advertisements were placed in newspapers, public service announcements were made on radio and television, and the study staff and physicians associated with the study were active on local radio and television shows. A nationally representative board of patrons, all distinguished women, helped immeasurably by participating in public service announcements and by generating publicity. An important and successful publicity measure was an insert with the May 1982 family allowance cheque, mailed by the Department of National Health and Welfare. Publicity was augmented when the then federal minister of health became a participant.

Nearly 100 000 letters were sent as group mailings to employees of large institutions and to members of professional associations.

Major efforts were made to inform physicians in each of the study areas about the study. A pamphlet was sent to physicians identified by the College of Family Physicians of Canada (Ontario). Recruitment through family physicians was not attempted, because pilot studies had shown that this approach was not likely to be successful. We identified each participant's family physician (to whom we sent reports on the screening results), but their approval for the recruitment of their patients was not sought.

Because of the extensive use of these overlapping mechanisms to recruit participants, it is difficult to determine their relative success. However, women were asked to record their sources of awareness of the study on entry.[22] In brief, they heard about the study mainly through the newspaper (36.9%), friends (32.1%) and the radio (30.2%). Only 6.6% heard about the study from their physician.

Randomization

The eligibility criteria, described in detail elsewhere,[1] were age 40 to 49 years, no mammography in the previous 12 months, no history of breast cancer, no current pregnancy and a signed consent form.

Randomization was individual and stratified by centre and 5-year age group. It was done by the local coordinator, who used lists supplied by the central office with preprinted identification numbers and group designations. To equalize group assignments after every two pages and to avoid large runs of assignments to one group, randomization was done in blocks. The blocking of the lists was unknown to the staff at the screening centres. Two sets of computer-generated random numbers, one blocked in twos, the other in fours, were used. An underlying purpose was the efficient use of the mammography equipment.

After the consent form was signed and the examination by the nurse (or physician in the centres in Quebec) completed, the record was given to the local coordinator, who wrote the participant's name on the next available line on one of the lists (according to 5-year age group). This procedure also assigned the woman an identification number. Assignment occurred regardless of whether the examination revealed an abnormality. The examiners did not have access to the lists.

Data forms were shipped to the central office once a month along with the original randomization sheets and a copy of the current pages still in use. These sheets were checked for violations, such as overwritten name spaces, erasures and unused slots.

Intervention

Women were randomly assigned to undergo either annual mammography screening and physical examination of the breasts (MP group) or a single physical examination of the breasts and annual follow-up through a mailed, self-administered questionnaire (UC [usual care] group). The first 62% of the women who entered the study were eligible for a 4-year program; the remainder were offered a 3-year program.

All women were taught breast self-examination, because it is regarded as part of the standard baseline care for breast cancer detection by the Canadian Cancer Society. For those who returned to the screening centres for examination, teaching of breast self-examination was reinforced.

Procedures

In 12 of the 15 study centres nurses provided the physical examination of the breasts; in the remaining 3 centres, in Quebec, physicians were required to perform this task. If during the physical examination of the breasts or the mammography an abnormality was detected, the participant was referred to the NBSS review clinic, where it was the study surgeon's role to discuss the mammography findings with the study radiologist who had read the films, to examine the participant and to decide whether further diagnostic procedures were indicated. Regardless of the group assignment, these procedures could be diagnostic mammography, fluid aspiration, open biopsy or, in women with impalpable abnormalities on mammography, needle localization biopsy.

Once the study surgeon had recommended a diagnostic procedure and the participant had been informed, the centre telephoned the participant's physician and sent a letter conveying the recommendation. It was the responsibility of the woman's physician to decide whether to accept the recommendation and to choose who would implement it. However, many women referred to the NBSS review clinic because of an abnormality discovered at screening (especially on physical examination in the early phases of the study) were deemed not to require a diagnostic procedure. Such women were reassured, and a letter was sent to their physician.

Follow-up

The coordinator for each centre was responsible for ascertaining whether the recommended diagnostic procedures had been carried out and for collecting reports of the surgical and pathological procedures from the institutions where they had been performed. Diagnostic procedures could be done independently of the screening process between screening visits or at any time after entry for women in the UC group. These events were identified during the course of the study through the annual questionnaires. As soon as it was reported that a biopsy had occurred the coordinator began the same chain of action as that for screen-generated biopsies. A designated NBSS reference pathologist reviewed the slides of removed tissue. If there was disagreement between the community pathologist and the reference pathologist the slides were reviewed by a panel of three to five NBSS pathologists. All diagnoses of breast cancer were histologically verified.

Records for women who moved out of a study area were transferred to the NBSS centre closest to the new home to encourage continued attendance. Otherwise, follow-up was by mail, as for the UC group.

After participants completed their screening schedule, direct follow-up stopped for those with no diagnosis of breast cancer. Information about new diagnoses of cancer was retrieved by linkage with provincial cancer registries (in Nova Scotia, Quebec, Ontario, Manitoba, Alberta and British Columbia) and occasionally on an ad hoc basis through health care professionals associated with the study. In Ontario and Quebec data from the cancer registries on the number of cases of breast cancer were available to the end of 1988; in Alberta and British Columbia the data were complete to the end of 1989, and in Manitoba and Nova Scotia to the end of 1990. All women known to have breast cancer are followed up annually by the NBSS central office with their surgeon or other contact physician.

Ascertainment of death

The death of participants was determined in several ways. Questionnaires mailed to participants during the NBSS screening program could be returned unanswered with "deceased" marked on them. For women known to have breast cancer routine annual follow-up through their physicians indicated whether death had occurred. Linkage with the Canadian Mortality Data Base (CMDB), Statistics Canada, was performed to identify women who had died as of Dec. 31, 1988. (The CMDB includes Canadians who died while residing in the United States.) Also, several deaths outside North America were noted through the routine follow-up procedures.

Verification of cause of death

Death certificates were obtained for participants who died in Canada. Investigative procedures were initiated for all deaths of women known to have had breast cancer, for those whose death certificate mentioned breast cancer and for those whose cause of death was reported as unknown, unknown primary cancer, lung cancer, colon cancer or liver cancer. All other causes of death were accepted as certified.

Women whose deaths necessitated investigation fell into one of the following three categories.

1. *Women known to have breast cancer, those whose death certificate mentions breast cancer and those whose death certificate mentions liver cancer or unknown primary cancer and whose clinical records refer in any way to a possibility of breast cancer.* The case file was reviewed by a panel comprising two of six available surgical, medical or radiation oncologists. The file included surgical and pathological reports associated with diagnosis and treatment, clinical records documenting the natural history of the disease and records describing the terminal course of the disease. All identifying information was removed, and the reviewers were blind to group assignment. According to the evidence provided, the reviewers decided whether they considered that the death was due, probably due, not due or probably not due to breast cancer. In rare cases of disagreement a third reviewer was enlisted. Since not all women with breast cancer died of breast cancer, the panel had to decide whether there was another underlying cause.

2. *Women whose cause of death was recorded as unknown or unknown primary cancer.* Investigations included acquisition of pathological reports and clinical notes, autopsy reports and coroner reports. In the coroner reports many of the causes of death were violent. In such instances or if the autopsy or pathological reports indicated an organ site clearly unrelated to the breasts there was no referral to panel review. However, if the cancer site appeared to be ambiguous a panel review was invoked.

3. *Women whose cause of death was recorded as lung or colon cancer.* Because of the possibility of misclassification in such cases, pathological reports, hospital records and autopsy reports were collected whenever available; in most cases the death certificate proved to be accurate. If the hospital records revealed disagreement among pathologists as to whether the primary site of cancer was the breast, slides were reviewed by an external panel of NBSS pathologists. If the breast was found to be involved the case was referred to panel review.

NBSS database

The database included records for 50 430 eligible women enrolled from January 1980 through March 1985. The coordinator at each screening centre checked the participant's questionnaire for completeness, especially the questions related to eligibility. All questionnaires and medical forms generated by the study were processed centrally. Extensive quality control was carried out during data collection. When forms were filled out incorrectly or critical fields were not filled in, the relevant screening centre was asked to identify and correct the errors.

Risk factors were recorded from the initial enrolment form and an epidemiologic questionnaire. The original version of the questionnaire has been published.[1] The version completed by most of the participants differed only in that it was commercially printed; the variables (all pretested in the pilot study in Toronto) remained the same.

In cases of breast cancer the reported size and grade of the tumour as well as the axillary node status were determined by pathologists in community hospitals who conformed to their own standards of practice. Because of this lack of uniformity, tumour grades were not incorporated into our analysis. For tumour size, available data were used, even though for mixed in-situ and invasive tumours the invasive component was not always measured. If the community pathologist noted a small microinvasive component but did not specify its size, such tumours were tabulated as being less than 5 mm in size. However, for several tumours no size was recorded.

Terminology

Screen 1 (*2, 3, 4* or *5*) was used to denote an event associated with the consecutive screening examinations. Thus, there can be rates of referral to the review clinic, biopsy and cancer detection for screen 1 (2, 3, 4 or 5). These events were independent of calendar year, since screen-1 exams could occur in any calendar year from 1980 to 1985. *Screen-detected cancer* was diagnosed as a direct result of a recommendation from the NBSS review clinic. *Interval cancer* occurred less than 12 months after a screening examination that did not result in a recommendation for diagnostic evaluation. (In this context interval cancer could occur in women with or without an abnormality suspected on either physical examination of the breasts or mammography. Thus, if an abnormality was suspected that led to a referral to the review clinic but no recommendation was made for a diagnostic procedure, the screen result was regarded as "negative" even if a cancer was diagnosed at the same site within the ensuing 12 months.) Cancers designated as interval 1 (2, 3, 4 or 5) reflected the previous screening examination. *Incident cancer* occurred more than 12 months after the previous screening examination.

Analysis

Student's *t*-test was used to determine the significance of differences in proportions between the two groups. A two-sided α level of 0.05 was used as the cutoff for statistical significance. Only those values of less than 0.05 are cited in the text. For all ratios of observed to expected cases of and deaths from breast cancer 95% confidence intervals (CIs) were computed.

Death due or probably due to breast cancer was the main end point for analyses. Death rates were computed with a person-year analysis on the basis of stratification by 5-year age group and by centre; it was assumed that all women not known to be dead were alive.

The observed numbers of deaths from breast cancer were compared with those expected (from national data) with the use of a program described by Morrison, Brisson and Khalid:[23] the time of normal detection of breast cancer in women whose age distribution and length of follow-up were similar to those of the participants was calculated. Population-based case-fatality rates were applied to estimate the expected number of deaths from breast cancer. Canadian data on breast cancer incidence were substituted for the US data used by Morrison and associates.

Life-table analysis was used to determine survival rates among women with invasive breast cancer.[24] Deaths from causes other than breast cancer were withdrawn from the life table. For comparisons of survival within groups the date of diagnosis was used as the entry point in the life table. Survival probabilities were compared with the use of the log-rank test.[25]

Exclusion from the analysis

Of the 50 472 women entered into the study 42, distributed equally between the two groups, were excluded from the analysis for the following reasons: (a) total refusal (4 participants withdrew from the study after group assignment and demanded to have their study records destroyed), (b) wrong screening procedure (25 women did not undergo mammography and should have or vice versa), (c) wrong age (10 women were less than 40 years at entry) and (d) recent mammography (3 participants had undergone mammography within the year before entering the study).

The protocol violations judged not to require exclusion were as follows: (a) double assignment (in 22 cases two women were given the same identification number because the coordinator failed to enter a woman's name on the list after assigning her to a group and then entered another name on that line; the error was caught at the time of data entry, and the woman first given the number was assigned another one because her name was not on the list) and (b) wrong age list (in 178 cases women were assigned using the wrong age list because of an error in calculating their age from the birth date on the questionnaire).

Presentation of data

Age was defined as age at entry. Attained age was not used in any analysis. Data were presented as numbers, rates per 1000 and proportions, as appropriate. For economy of presentation the numerators and denominators from which the rates and proportions were derived have not been included in some cases. These data are available upon request from the authors.

Results

The active recruitment phase commenced in Toronto in January 1980, Quebec in August 1980, Montreal, Hamilton and Winnipeg in 1981, most of the remaining centres in 1983 and Alberta in 1984. The half-way point of enrolment was reached in July 1983. Recruitment stopped in the first five centres in 1984 and in the remainder at the end of March 1985. Of the 50 430 women included in the analysis 6% were recruited in 1980, 14% in 1981, 19% in 1982, 25% in 1983, 32% in 1984 and 3% in 1985. Screening continued until June 1988. The follow-up period ranged from 5.4 to 12 (mean 8.5) years.

Characteristics of the study population

Detailed analyses of the epidemiologic characteristics reported on the questionnaire were performed by centre and province. An analysis by single year of age indicated almost an equal distribution between the two groups. The number of women entered fell off with increasing age within each

5-year category, an indication that the younger women were recruited more successfully.

Table 1 summarizes the data for other epidemiologic variables by group. The last column gives the data for the Canadian population matched for age and sex.[26] Differences between the two groups were minimal, less than 1% for a given characteristic in most instances. In addition to the data presented in Table 1, there were only minimal differences in oral estrogen use (mean duration 5.7 years).

Compared with the Canadian population the participants had a similar distribution of marital status but differed in other respects. They were less likely to have more than three children and more likely to have one or two. Substantially more of them had trade or business training or a university education. More had been born in North America and fewer in Europe or elsewhere. Slightly fewer had never smoked, and even fewer were heavy smokers. The average cigarette consumption was 17 per day over an average duration of 17.2 years (not shown in Table 1). More worked in health-related, teaching, managerial or administrative, science-related or technology-related occupations, with correspondingly fewer in sales and service and "other" occupations.

Compliance with screening

In the MP group full compliance with screening after screen 1 (when, by definition, compliance was 100%) varied from 85.6% (for screen 5) to 89.4% (for screen 2). In addition, a small proportion (1.7% to 2.9%) of the women accepted physical examination but refused to undergo mammography. Of the women in the MP group 3.5% to 6.5% missed one or more screens after screen 1 but still submitted questionnaires. Over 90% of the participants in the UC group (from 93.3% to 94.9% in the various years) returned their annual questionnaire.

Referral to review clinic

Table 2 displays the reasons for referral to the NBSS review clinic. Referrals were more frequent in the MP group than in the UC group because of mammographic abnormalities detected in the absence of physical findings. The contribution of physical findings to the referral rate was almost equal in the two groups at screen 1 (14.1% in the MP group and 14.6% in the UC group).

Impact of recommendations from the screening centres

Table 3 gives the diagnostic procedures recommended by the study surgeons and the procedures actually performed, since community surgeons did not necessarily agree with the procedures recommended. In general, more procedures were recommended and performed in the MP group than in the UC group, and more were performed at screen 1 than at subsequent screens. Diagnostic mammography in the community was sometimes recommended by the study surgeon, more often for women in the UC group than for those in the MP group. More diagnostic mammography was performed than recommended; this reflected a need in the community for mammography before proceeding to biopsy as recommended by the study surgeon. All of the screening centres provided the mammograms on request to the community physician or institution to which the participant was referred. Hence many of the NBSS mammograms were used in the community for diagnostic purposes.

Mammography was also performed in the community and was reported by participants on the annual questionnaires. In some cases the mammography was requested to investigate abnormalities. The numbers of women reporting one or more community mammograms during the study period were 1790 (7.1%) of those in the MP group and 6651 (26.4%) of those in the UC group. The proportion in the MP group remained stable across the screening years, ranging between 2.3% and 2.6%. In contrast, the proportion increased over time in the UC group, from 7.0% between years 1 and 2 to 18.1% between years 4 and 5.

Table 4 shows the benign biopsy rates; only surgical biopsies, with or without needle localization, were included. The relatively high rates reflect the large numbers of biopsies done in community institutions as the definitive diagnostic test. The rates were particularly high in the MP group and at screen 1.

Cancer detection rates

The rates of screen-detected cancer, including in-situ and invasive cancer, are shown in Table 5 by year of screening examination. Nonattenders were not included in the denominator for screens 2 through 5. In the MP group the rates were higher at screen 1 than at other times. At screen 1 the rate of detection by physical examination, alone or in combination with mammography, was higher in the MP group than in the UC group (2.70 v. 2.46 per 1000); this difference was not statistically significant.

The rates of interval cancer are presented in Table 5. The denominator was the number of women in the same group who had attended the previous screen. The rate for interval 1 was higher in the UC group than in the MP group, but the

Table 1: Demographic characteristics of women aged 40 to 49 years upon entry into the Canadian National Breast Screening Study (NBSS) and women in the general population

Characteristic	Study group;* no. (and %) of women MP group (n = 25 214)	UC group (n = 25 216)	% of women in Canada
Marital status	(n = 25 170)	(n = 25 171)	
Never married	1 639 (6.5)	1 636 (6.5)	6.1
Married	20 296 (80.6)	20 321 (80.7)	81.2
Separated or divorced	2 679 (10.6)	2 706 (10.8)	9.6
Widowed	556 (2.2)	508 (2.0)	3.1
No. of live births†	(n = 23 472)	(n = 23 459)	
0	2 331 (9.9)	2 369 (10.1)	7.6
1	2 353 (10.0)	2 447 (10.4)	9.7
2	7 943 (33.8)	7 774 (33.1)	27.2
3	6 281 (26.8)	6 186 (26.4)	24.2
4	2 848 (12.1)	3 008 (12.8)	15.4
5	1 032 (4.4)	1 072 (4.6)	7.6
≥ 6	684 (2.9)	603 (2.6)	8.1
Reproductive status	(n = 25 214)	(n = 25 216)	
Premenopausal	16 739 (66.4)	16 922 (67.1)	–
Perimenopausal	298 (1.2)	298 (1.2)	–
Postmenopausal	1 239 (4.9)	1 204 (4.8)	–
Underwent hysterectomy and oophorectomy	1 586 (6.3)	1 492 (5.9)	–
Underwent hysterectomy	4 873 (19.3)	4 834 (19.2)	–
Underwent oophorectomy	171 (0.7)	165 (0.7)	–
Unknown	308 (1.2)	301 (1.2)	–
Level of education	(n = 23 001)	(n = 22 926)	
Grade 8	1 892 (8.2)	1 956 (8.5)	25.1
Grade 9–13	7 011 (30.5)	7 026 (30.6)	40.4
Trade or business school	8 933 (38.8)	8 735 (38.1)	22.3
University	5 165 (22.5)	5 209 (22.7)	12.2
Family history of breast cancer, family member	(n = 9 493)	(n = 9 652)	
Mother	2 051 (8.1)	2 055 (8.1)	–
Sister	831 (3.3)	872 (3.5)	–
Daughter	2 (0.0)	4 (0.0)	–
Second-degree relative‡	6 609 (26.2)	6 721 (26.7)	–
Place of birth	(n = 25 214)	(n = 25 216)	
North America	21 246 (84.3)	21 266 (84.3)	76.1
Europe	3 325 (13.2)	3 284 (13.0)	18.1
Elsewhere	601 (2.4)	645 (2.6)	5.8
Not available	42 (0.2)	21 (0.1)	–
Cigarette smoking status	(n = 25 214)	(n = 25 216)	
Never smoked	12 074 (47.9)	12 034 (47.7)	53.9
Smoked, no. of cigarettes			
1–10	1 968 (7.8)	1 931 (7.7)	7.8
11–20	2 355 (9.3)	2 351 (9.3)	13.0
> 20	2 306 (9.1)	2 249 (8.9)	14.0
Used to smoke	6 511 (25.8)	6 651 (26.4)	11.4
Occupation	(n = 23 905)	(n = 23 922)	
Not in workforce§	7 912 (33.1)	7 874 (32.9)	32.0
Clerical	5 200 (21.8)	5 289 (22.1)	22.4
Medical or health related	2 615 (10.9)	2 574 (10.8)	5.4
Teaching	2 148 (9.0)	2 156 (9.0)	4.7
Managerial or administrative	1 824 (7.6)	1 807 (7.6)	4.3
Science or technology related	768 (3.2)	781 (3.3)	1.6
Sales, service	2 162 (9.0)	2 210 (9.2)	17.6
Other	1 276 (5.3)	1 231 (5.1)	12.0

*MP = mammography and physical examination (PE) of the breasts, UC = usual care (single PE of the breasts and annual follow-up by mailed, self-administered questionnaire).
†Single women not included for comparability with the Canadian population.
‡Includes aunts, cousins and other relatives.
§Includes women who were housewives, retired or unemployed.

difference was not statistically significant. In the MP group the rate was highest for interval 1.

The denominator for the rate of incident cancer was the number of women in the same group who had not attended a screen for over 12 months (Table 5). For women eligible to return for screening at screens 2 through 4 (those in the MP group) the denominator was the number of women who had not returned at the visit before detection. Since one-third of the study population was eligible for four screens

Table 2: Frequency of referral to review clinic, by screen

Group; method by which abnormality detected	Screen; no. (and %) of women 1	2	3	4	5
MP group	(n = 25 214)	(n = 22 424)	(n = 22 066)	(n = 21 839)	(n = 14 146‡)
Mammography (Ma) only	1 118 (4.4)	557 (2.5)	531 (2.4)	489 (2.2)	306 (2.2)
PE only	3 137 (12.4)	1 505 (6.7)	1 216 (5.5)	1 115 (5.1)	681 (4.8)
Ma and PE	432 (1.7)	135 (0.6)	126 (0.6)	86 (0.4)	69 (0.5)
All PE*	3 569 (14.1)	1 640 (7.3)	1 342 (6.1)	1 201 (5.5)	750 (5.3)
UC group†	(n = 25 216)				
PE	3 674 (14.6)	–	–	–	–

*Number of women whose abnormality was detected by PE alone or in combination with Ma.
†No figures are given for screens 2 through 5 because women in the UC group were not eligible for rescreening.
‡Only 62% of the women were eligible for the fifth screen.

Table 3: Diagnostic procedures recommended* (R) and performed† (P) per 1000 women, by screen

	Screen									
	1		2		3		4		5	
Procedure; group	R	P	R	P	R	P	R	P	R	P
Fluid aspiration										
MP	19.2	13.2	13.7	11.7	13.1	12.2	12.9	11.4	14.4	13.4
UC	15.3	12.5	–	–	–	–	–	–	–	–
Tissue aspiration or needle biopsy										
MP	16.6	13.4	10.1	8.6	9.4	7.9	7.5	6.1	7.0	5.9
UC	14.4	11.1	–	–	–	–	–	–	–	–
Open surgical biopsy										
MP	26.5	25.3	13.2	11.9	10.8	9.8	9.4	7.6	9.0	8.1
UC	16.8	13.4	–	–	–	–	–	–	–	–
Needle localization biopsy										
MP	16.2	12.4	8.9	7.6	7.7	5.5	8.8	7.0	10.2	7.4
UC	0.4	0.3	–	–	–	–	–	–	–	–
Diagnostic Ma										
MP	1.1	7.7	0.4	1.2	0.3	1.2	0.4	1.5	0.6	1.4
UC	14.6	15.1	–	–	–	–	–	–	–	–

*Procedures recommended by an NBSS surgeon.
†Procedures performed in the community.

Table 4: Rates of biopsy detection of benign lesions per 1000 women, by screen

	Screen; biopsy rate									
	1		2		3		4		5	
Screening method	MP	UC	MP	UC	MP	UC	MP	UC	MP	UC
Ma only	18.1	–	10.2	–	7.9	–	7.6	–	8.8	–
PE only	9.9	11.5	5.8	–	4.0	–	3.6	–	3.4	–
Ma and PE	5.6	–	1.6	–	1.3	–	0.8	–	1.0	–
Overall rate	33.6	11.5	17.5	–	13.1	–	12.0	–	13.2	–
n	25 214	25 216	22 424	–	22 066	–	21 839	–	14 146	–

only, it became the denominator in the fifth year after entry. For women not eligible to return for screening after screen 1 (those in the UC group) the denominator was the number of women who were still alive at the relevant time.

In Table 5 the incident-4 rate in the MP group, although high, was based on small numbers. The women in the UC group had relatively stable rates of incident cancer. For women who entered the NBSS in 1984 and 1985 the data were probably incomplete at 6 and 7 years after entry; therefore, the rates of incident cancer for these years were not included in Table 5 (the numbers of women with breast cancer ascertained to date in these years are 47 in the MP group and 41 in the UC group).

The observed rates of invasive cancer for the 5 years after entry were compared with the expected rates (determined on the basis of data from Statistics Canada for 1980–86) (Table 6). Only cases of invasive cancer were included because cancer registries do not always have data for cases of in-situ cancer. The numerators were women with screen-detected interval and incident invasive cancer. The denominators were women-years at risk. Year 1, the 12-month period from the date of entry, included women whose cancer was detected at screen 1 and in interval 1. Year 2 included those whose cancer was detected at screen 2 or in interval 2 or was classified as incident 2. The cumulative rates for years 2 through 5 are at the bottom of Table 6. In the MP group the cumulative ratio of observed to expected rates of breast cancer was 1.26 (95% CI 0.98 to 1.59); however, this was not statistically significant. In the UC group the cumulative ratio was 1.02 (95% CI 0.78 to 1.33).

Nodal status and tumour size of invasive cancers

At screen 1 more node-negative tumours and tumours with four or more nodes involved were detected in the MP group than in the UC group (Table 7). The numbers of incident cancers observed among the women in the UC group for years 2 through 5 were high because these were the women screened only once, on entry.

Table 7 presents the nodal status by mode of

Table 5: Detection rates of breast cancer, including in-situ cancer, per 1000 women, by year

Type of cancer; screening method	Year; detection rate									
	1		2		3		4		5	
	MP	UC	MP	UC	MP	UC	MP	UC	MP	UC
Screen-detected cancers										
Ma only	1.19	–	0.76	–	0.77	–	1.37	–	0.78	–
PE only	1.35	2.46	0.62	–	0.54	–	0.50	–	0.71	–
Ma and PE	1.35	–	0.36	–	0.68	–	0.50	–	0.35	–
Overall rate	3.89	2.46	1.74	–	1.99	–	2.38	–	1.84	–
n	25 214	25 216	22 424	–	22 066	–	21 839	–	14 146	–
Interval cancers										
Overall rate	0.75	1.11	0.71	–	0.36	–	0.46	–	0.64	–
n	25 214	25 216	22 424	–	22 066	–	21 839	–	14 146	–
Incident cancers										
Overall rate	–	–	0.76	1.51	1.03	1.68	2.94	1.84	0.61	1.57
n	–	–	2 645	25 092	2 917	25 033	3 057	24 954	24 792	24 883

Table 6: Observed and expected* incidence rates of invasive breast cancer and cumulative rates per 1000 women, by year

Year	MP group				UC group			
	No. of person-years	Observed rate	Expected rate	Ratio	No. of person-years	Observed rate	Expected rate	Ratio
1	25 085	103	31.7	3.2	25 113	83	31.7	2.6
2	25 009	41	33.5	1.2	25 055	35	33.6	1.0
3	24 928	49	35.3	1.4	24 994	37	35.4	1.0
4	24 834	52	37.0	1.4	24 918	39	37.2	1.0
5	24 765	39	38.8	1.0	24 863	37	39.0	0.9
Cumulative rates†		72.7	57.8	1.26		59.3	57.9	1.02
95% confidence interval (CI)				0.98–1.59				0.78–1.33

*Expected rates were based on data from Statistics Canada, 1980–1986.
†Rates for years 2 through 5.

detection for screen-detected invasive cancers. The mode of detection in the UC group was physical examination only, but in the MP group it was mammography only or physical examination (alone or in combination with mammography). At screen 1 there were more cases of cancer with four or more nodes involved detected by means of physical examination in the MP group than in the UC group (17 v. 5) ($p < 0.017$). For all screening years combined, 67 (33%) of the 205 cases of invasive cancer in the MP group were detected by mammography alone; 53 (79%) of them were known to be node negative. Of the 138 detected by physical examination (alone or in combination with mammography) 81 (59%) were known to be node negative.

An analysis of the recorded tumour size for all cases of invasive cancer has been performed in the same detail as for nodal status (the findings are available upon request from the authors). As mentioned previously, tumour size was recorded by different community pathologists with no possibility for a uniform approach. At screen 1, 44 (51%) of the 86 invasive tumours detected in the MP group and 26 (43%) of the 60 detected in the UC group were reported to be small (less than 20 mm in diameter). In general, such tumours accounted for at least 38% of all incident or interval cancers in the two groups; for cases in the MP group detected in interval 1 and classified as incident 2 through 5 cancers the proportions were 25% and 32% respectively.

For all years combined, the MP group had more small tumours than the UC group if screen-detected, interval and incident tumours are combined (156 v. 116). Also, the number of large tumours (20 mm or more) was slightly higher in the former group (119 v. 101).

Mortality results

Table 8 provides the causes of death ascertained from the CMDB, unless the review panel reclassified the cause. There were several differences between the two groups in the number of deaths by cause. Compared with the UC group, the MP group had more women who died of breast cancer, ovarian cancer, pancreas cancer, hematopoietic neoplasms, central nervous system disorders (nonvascular) and circulatory problems. The reverse was true for women who died of colorectal cancer, other types of cancer and miscellaneous causes. Only the difference in the number of deaths from colorectal cancer was significant ($p < 0.014$). The total number of deaths was almost equal in the two groups.

Table 9 presents the number of deaths from

Table 7: Nodal status of all invasive tumours, by year*

	Screen-detected cancer							
	MP				Interval cancer		Incident cancer	
Year; no. of nodes involved	All	Ma alone	PE†	UC	MP	UC	MP	UC
Year 1								
None	52	17	35	34	7	13	–	–
1–3	14	1	13	16	2	5	–	–
≥ 4	19	2	17	5	6	2	–	–
Unknown	1	1	0	5	1	4	–	–
Total	86	21	65	60	16	24	–	–
Years 2–5								
None	82	36	46	–	20	–	10	92
1–3	24	7	17	–	8	–	4	20
≥ 4	9	1	8	–	5	–	5	15
Unknown	4	2	2	–	4	–	9	21
Total	119	46	73	–	37	–	28	148
Year 6 or more								
None	–	–	–	–	–	–	20	18
1–3	–	–	–	–	–	–	3	2
≥ 4	–	–	–	–	–	–	3	1
Unknown	–	–	–	–	–	–	19	19
Total	–	–	–	–	–	–	45	40
All years								
None	134	53	81	34	27	13	30	110
1–3	38	8	30	16	10	5	7	22
≥ 4	28	3	25	5	11	2	8	16
Unknown	5	3	2	5	5	4	28	40
Total	205	67	138	60	53	24	73	188

*For all dashes there was no cancer in this category by study design.
†Tumours detected at physical examination, alone or in combination with Ma.

breast cancer 7 years after entry. (Fewer deaths from breast cancer were recorded in Table 8 because the cutoff for the linkage with the CMDB was Dec. 31, 1988.) For screen 1 there were more deaths in the MP group than in the UC group. There were no deaths in the UC group in the categories for screens 2 to 5 or intervals 2 to 5 because there was no screening in this group after screen 1. The deaths in the UC group that correspond to these categories in the MP group largely appear as deaths among the women with incident cancer. The total numbers of deaths from breast cancer were 38 in the MP group and 28 in the UC group; the difference was not significant. The ratio of the proportions of death from breast cancer in the MP group compared with those in the UC group was 1.36 (95% CI 0.84 to 2.21).

The lower part of Table 9 displays the observed and expected cumulative rates of death from breast cancer at 7 years after entry. In the MP group the observed rate was slightly higher than expected, whereas in the UC group it was lower than expected.

Table 8: Causes of death to end of 1988, by study group

Cause*	Group; no. (and %) of women	
	MP group	UC group
Cancer		
Breast	29 (18.2)	18 (11.5)
Colorectal	5 (3.1)	18 (11.5)
Hematopoietic	11 (6.9)	6 (3.8)
Lung	19 (11.9)	18 (11.5)
Ovarian	12 (7.5)	8 (5.1)
Other gynecologic	0	3 (1.9)
Pancreas	10 (6.3)	6 (3.8)
Stomach	3 (1.9)	5 (3.2)
Other	16 (10.1)	23 (14.7)
Central nervous system disorder (nonvascular)	5 (3.1)	1 (0.6)
Circulatory disorder	21 (13.2)	16 (10.3)
Endocrine or metabolic condition	1 (0.6)	3 (1.9)
External cause (violent)	19 (11.9)	18 (11.5)
Infection or parasitic disease	2 (1.3)	0
Miscellaneous	3 (1.9)	8 (5.1)
Respiratory disease	2 (1.3)	4 (2.6)
Unknown	1 (0.6)	1 (0.6)
Total	159	156

*The tabulated underlying cause of death was calculated on the basis of all available information.

Table 9: Cumulative number of deaths from breast cancer 7 years after entry, by study group and time of breast cancer detection

Time of detection	Group; no. of deaths	
	MP	UC
Screen 1	11	5
Screens 2–5	11	–
Interval 1	5	5
Intervals 2–5	4	–
Incident (> 12 mo after last screen)		
Among noncompliers	5	–
After scheduled end of screening	2	18
Total	38	28
Cumulative rates*		
Observed	14.7	10.4
Expected	13.7	13.7
Ratio	1.08	0.76
95% CI	0.59–1.80	0.36–1.40

*Rates are per 10 000 person-years.

However, the differences between the observed and expected rates were not statistically significant.

The survival rates from the time of entry were high and similar in the MP and UC groups: at 7 years 90.2% and 89.9% respectively of the women with invasive cancer were alive. Table 10 displays the survival rates from the date of diagnosis. Again, data for the UC group were only available after screen 1 as incident cancers. Thus, comparisons of the survival rates between the two groups for interval and incident cancers were based on cancers diagnosed at different times from entry and are not valid. The only valid comparison of observed survival rates is for cases detected at screen 1: the survival rate was better in the UC group than in the MP group, but not significantly so. Table 10 provides the survival rates by mode of detection in the MP group. The women whose cancer was detected by mammography only had the highest survival rate, but it was not significantly higher than that among women whose cancer was detected by physical examination (alone or in combination with mammography).

Discussion

Study population and procedures

The NBSS is the largest randomized study reported to date evaluating breast screening among women aged 40 to 49 years. The close matching of the demographic and personal characteristics of the women in the two study groups overall and by centre confirms the validity of the group assignment procedures.

The participants differed in several important respects from the Canadian population. In particular, they had a higher socioeconomic status. This and other analyses[27-29] have shown that the participants had, if anything, more risk factors for breast cancer; therefore, their expected incidence rate of breast cancer should be at least as great as the rate in the Canadian population, a finding similar to that in other screening studies.[30]

Since those who assigned the women to the two groups used lists with the randomization pre-entered, one must consider whether bias occurred. If there had been a bias to assigning women with symptoms, obvious risk factors (e.g., a positive family history) or a palpable mass to the MP group we would have detected inequalities in the distribution of relevant risk factors; no such evidence was found. Furthermore, such bias for those with a palpable mass was unnecessary, because all women with abnormalities detected at the physical examination were to be referred to the study surgeon. Our quality control included checks on the randomiza-

Table 10: Survival rates among women with breast cancer from date of diagnosis, by detection category and study group

Year from diagnosis	Screen 1		Screens 2–5		Screen-detected cancer (MP)		
	MP	UC	MP	UC	Ma alone	PE alone	Ma + PE
1	100.0	100.0	100.0	–	100.0	100.0	100.0
2	100.0	100.0	99.4	–	100.0	100.0	98.6
3	95.9	98.4	98.8	–	99.0	96.2	97.3
4	91.8	98.4	93.3	–	95.0	90.8	91.6
5	90.7	93.3	92.5	–	95.0	89.3	90.0
6	88.7	93.3	92.5	–	95.0	89.3	86.4
7	88.7	91.2	92.5	–	95.0	89.3	86.4
No. at time zero	98	62	161	–	105	81	73
No. (and %) of deaths	11 (11.2)	5 (8.1)	11 (6.8)	–	5 (4.8)	8 (9.9)	9 (12.3)

Year from diagnosis	Interval cancers		Incident cancers	
	MP	UC	MP	UC
1	100.0	100.0	100.0	100.0
2	94.9	92.9	96.5	97.8
3	93.2	89.3	93.8	97.1
4	91.4	89.3	87.3	92.1
5	87.4	85.6	82.9	91.3
6	82.1	81.8	82.9	88.3
7	82.1	81.8	82.9	88.3
No. at time zero	62	28	76	206
No. (and %) of deaths	9 (14.5)	5 (17.9)	6 (7.9)	16 (7.8)

tion process. Errors or falsifications that could have led to bias were not detected. The original randomization sheets were carefully rechecked, specifically in relation to women who died; no evidence of any falsification, erasure or other changes was found. The proportion of women referred for review at screen 1 on the basis of findings at physical examination was similar in the two groups (Table 2). Therefore, assignment bias did not occur.

The NBSS is unique among randomized trials of breast screening published to date because it collected data on risk factors for breast cancer from all of the participants. For a substantial proportion of the women self-administered dietary questionnaires were also completed and so far have been used in two studies.[31,32] Because of the lack of data on risk factors in recent trials (the HIP study evaluated risk factors only in samples of control subjects and women who refused to enter the study[33]), comparison is not possible between the NBSS participants and those in other trials.

Cancer detection

The cancer detection rates in the MP group (Table 5) correspond to those reported from previous studies.[34,35] As expected, the rates were higher in the MP group than in the UC group at screen 1. These rates were achieved at the cost of substantial benign biopsy rates, especially for abnormalities detected by mammography only (Table 4); this reflected the fact that the screening centres could only make recommendations to the participants' family physicians and the main diagnostic approach in the community hospitals was biopsy. For subsequent screens the benign biopsy rates were lower. The benign to malignant biopsy ratios were substantially higher than those reported from European studies.[36,37]

In assessing our interval cancer rates (Table 5) one must recognize that we included women who had had an abnormal finding on the previous screen, either through mammography or physical examination, but for whom a recommendation for further evaluation was not made by the study surgeon. Thus, our rates cannot be compared with those reported from other studies that involved only women with negative findings at the previous screen. Furthermore, all of the NBSS participants were taught and urged to practise breast self-examination, which led to the detection of cancer in some cases, thus potentially reducing the delay in diagnosis of breast cancer.

Substantially more node-negative and small tumours were detected by mammography, and there was a lower interval-1 rate in the MP group than in the UC group. Yet, with up to 7 years of follow-up from entry, there was no reduction in the cumulative incidence of large or node-positive tumours at the time of diagnosis in the MP group compared with the UC group. One possible explanation could be underascertainment of disease in the UC group, but there was little evidence of this, since the ratio of observed to expected cases of breast cancer during years 2 through 5 was 1.02 (Table 6).

The cumulative number of women with "advanced" cancer was higher in the MP group than in the UC group; this appears to have been because more cases of advanced cancer were detected at screen 1 in the MP group. An analysis of nodal status, in which the screening centres were classified into one of three groups depending on their technical proficiency (low, medium or high), as determined by the scores from the external review of mammograms,[21] revealed that this excess was not associated with mammography "quality." Thus, of the tumours with four or more nodes involved that were detected at screen 1, six in the MP group and two in the UC group were detected by centres with a low level of proficiency, eight and two respectively by those with a medium level, and five and one respectively by those with a high level. Cross-classification of tumour size and nodal status for screen-1 detections (not tabulated here) showed that much of the excess in the MP group at screen 1 was due to small tumours with four or more nodes involved. Such tumours may simply have been missed at review in the UC group; the availability of mammography in the MP group for all women referred because of abnormalities at physical examination and the required consultation between a study surgeon and a study radiologist at or just before the time of review probably led to a greater efficiency in cancer detection in the MP group than in the UC group. This is supported by the higher detection rates associated with physical examination findings at screen 1 among women in the MP group than among those in the other group (Table 5). Although diagnostic mammography was available at the time of review in the UC group, it was performed in only 5.4% of such women.

Previous studies have shown a small excess in the number of advanced tumours or deaths in the screened group compared with the control group among women under the age of 50 in the early years of follow-up[38,39] and among women under 55 in the first 5 years of follow-up.[40] Our excess of advanced tumours may be more pronounced because we enrolled the largest number of women aged 40 to 49, all through screening centres. In other studies invitations to attend screening were randomized, and the women who refused, not possible in the NBSS, could have had a diluting effect on the detection of an excess of advanced cancer.

Rates of death and survival

The death and survival rates were compatible with the cancer detection rates in that despite the early detection of both advanced and early cancers in the MP group, the rate of death from breast cancer did not decrease. The use of mammography and physical examination achieved the expected breast cancer detection rates, but the only survival advantage was for the women with breast cancer detected by mammography alone, presumably because of lead time.

Although these results were not expected when the NBSS was started they are less surprising now, because more women under 55 years of age died of breast cancer in the mammography group than in the control group in the first 5 years of the Malmö trial,[40] and a similar trend, albeit less evident, was seen among women less than 50 in the Nijmegen case-control study,[39] the Swedish two-county trial[41] and the Stockholm trial.[42] Other studies have shown no reduction in the death rate among women aged 40 to 49 on entry who were screened,[43,44] at least in the first 10 years after screening had been initiated.[6] In these studies, noncompliers contributed greatly to the preliminary death rates; such women did not exist at the initial screen in our study. Furthermore, other studies compared screening with no screening, whereas in the NBSS, women with clinically detectable breast tumours were identified in the two groups on enrolment but were not excluded.

Potential biases that might have influenced our results must be considered. As demonstrated earlier, randomization bias did not occur. The number of deaths from breast cancer since 1988, ascertained through follow-up of cases and through the cancer registry linkages, may not represent all such deaths. Nevertheless, the numbers were similar in the MP and UC groups (9 and 10 respectively). Given the almost complete absence of death from breast cancer in the early years after diagnosis, any cases of breast cancer diagnosed after the registry linkage cutoff dates are unlikely to have resulted in deaths that would materially affect the findings. Furthermore, the registry linkages covered 91% of the Canadian female population: any out-of-province mobility recorded during the follow-up through the screening centres had been incorporated in the files used for linkage. Thus, it seems unlikely that many deaths were missed or that the missed deaths would markedly change our observed distribution of deaths from breast cancer.

Since the total number of deaths as of 1988 was equal in the two groups, the greater number of deaths from breast cancer in the MP group and the greater number of deaths from colorectal cancer and other causes in the UC group raise the possibility of errors in death certification. The review of all deaths from lung cancer and colon cancer and the panel review process have diminished the likelihood of this. However, it could be that women with early detection of breast cancer in the MP group were deprived of appropriate investigative procedures when metastases were recognized and that these metastases were wrongly attributed to breast cancer by clinicians and the panel. The converse in the UC group would be to erroneously ascribe metastases to another primary site instead of breast cancer. Again, these possibilities are remote.

The main reason for the small difference in the death rate between the two groups was the higher number of tumours with four or more nodes involved detected at screen 1 in the MP group than in the UC group. Although many of these were relatively small tumours, which suggests that the mammograms available for interpretation of abnormalities detected at physical examination could have moved up the time of their diagnosis, the difference has not disappeared and cannot be fully explained.

Finally, the possibility that differences in treatment could have affected the observed death rates has been investigated,[45] since only effective treatment for screen-detected cancers can result in a reduced rate. The records of all the participants who died of breast cancer as of January 1990 and those of a sample of living participants with breast cancer matched for cancer stage and study centre were carefully reviewed. A medical and a radiation oncologist, blind as to group assignment and outcome, reviewed the details of chemotherapy, hormone therapy, radiotherapy and surgery. In most cases the treatment was considered appropriate for the stage of disease; there was no difference between the women who had died and those who were alive. Perhaps the treatment was appropriate for the stage but not for the particular biologic characteristics of the tumour, which could not be identified at that time because the use of biologic markers to indicate adverse prognostic potential in patients with node-negative disease is only now becoming established.

In conclusion, our results confirm those from previously published reports on mammography screening alone:[6] there is no evidence that screening for breast cancer is effective among women aged 40 to 49 years, at least in the first 7 years after initiation of screening. Priority must be given to explaining the difference in effectiveness of screening between women in that age group and those 50 years or older.

Follow-up of the women enrolled in the NBSS continues, and we plan to report the 10-year results in about 3 years.

We thank the women who volunteered to participate in this trial and who willingly gave of their time and energy knowing it was an experiment. We also thank the staff of the coordinating and local study centres, particularly the national coordinator, Charlotte Turnbull, and the local coordinators: L'ainée Duncan (Halifax); Andrée Christen (Quebec); Lorraine Grégoire, Chantal Perret and Lise Simard (Montreal); Judy Snider (Ottawa); Pat Carpick, Luby Carr, Carrol Clarkson, Ann Hampson and Bette Johnston (Toronto); Joan Alvarez and Margaret Opie (Hamilton); Maureen Campbell and Donna Dean (London); Pat Edward (Winnipeg); Carol May (Edmonton); Kay Johnson (Red Deer, Alta.); Louise Watson (Calgary); and Margaret Fryer (Vancouver). We are extremely grateful to the nurses and radiographers, whose skills contributed to the validity of the screening tests, and the secretaries, telephone callers, coders, data-entry personnel and programmers who helped with the trial. We thank the volunteers who contributed their time and energy at many of the screening centres.

This study was supported by the Canadian Cancer Society, the Department of National Health and Welfare, the National Cancer Institute of Canada, the Alberta Heritage Fund for Cancer Research, the Manitoba Health Services Commission, the Medical Research Council of Canada, le Ministère de la Santé et des Services sociaux du Québec, the Nova Scotia Department of Health and the Ontario Ministry of Health. Dr. Miller was supported in part by a National Health Scientist Award from the Department of National Health and Welfare.

The reference radiologist for the NBSS was the late Dr. Douglas V. McFarlane, Toronto Western Hospital; the reference physicist was Dr. Martin Yaffe, Ontario Cancer Institute. The conduct of the NBSS was guided by the Policy Advisory Group, appointed by the Department of National Health and Welfare and the National Cancer Institute of Canada; in the last few years of the trial the group comprised Dr. Vivien Basco (chair), Dr. Carol Buck, Dr. Richard Margolese, Dr. Alan S. Morrison (deceased), Dr. Ross Prentice, Professor Samuel Shapiro, Dr. Edward A. Sickles and Dr. Donald Wigle. The administration and financial aspects of the trial were guided by the Management Committee: Mr. Douglas Barr (chair), Dr. James W. Meakin, Dr. Peter G. Scholefield, Dr. Lou Siminovitch and Dr. May Smith.

References

1. Miller AB, Howe GR, Wall C: The National Study of Breast Cancer Screening: protocol for a Canadian randomized controlled trial of screening for breast cancer in women. *Clin Invest Med* 1981; 4: 227-258
2. Leger JL, Naimark AP, Beique RA et al: Report of the "ad hoc" committee on mammography. *J Can Assoc Radiol* 1974; 25: 3-21
3. Miller AB: Preface. *Clin Invest Med* 1981; 4: 231-232
4. Miller AB: Is routine mammography screening appropriate for women 40-49 years of age? *Am J Prev Med* 1991; 7: 55-62
5. Report of the Working Group to Review the National Cancer Institute-American Cancer Society Breast Cancer Detection Demonstration Projects. *J Natl Cancer Inst* 1979; 62: 639-709
6. Shapiro S, Venet W, Strax P et al: *Periodic Screening for Breast Cancer. The Health Insurance Plan Project and Its Sequelae, 1963-1986*, Johns Hopkins U Pr, Baltimore, 1988
7. Chu KC, Smart CR, Tarone RE: Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. *J Natl Cancer Inst* 1988; 80: 1125-1132
8. Miller AB, Chamberlain J, Day NE et al: Report on a workshop of the UICC project on evaluation of screening for cancer. *Int J Cancer* 1990; 46: 761-769
9. Tabár L, Fagerberg G, Duffy SW et al: Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992; 30: 187-210
10. Baines CJ: Impediments to recruitment in the Canadian National Breast Screening Study: response and resolution. *Controlled Clin Trials* 1984; 5: 129-140
11. Baines CJ, To T: Changes in breast self-examination behavior achieved by 89,835 participants in the Canadian National Breast Screening Study. *Cancer* 1990; 66: 570-576
12. Baines CJ, Wall C, Miller AB et al: Sensitivity and specificity for first screen mammography in the Canadian National Breast Screening Study. *Radiology* 1986; 160: 295-298
13. Baines CJ, McFarlane DV, Miller AB et al: Sensitivity and specificity of first screen mammography in 15 NBSS centres. *J Can Assoc Radiol* 1988; 39: 273-276
14. Baines CJ, Miller AB, Bassett AA et al: Physical examination: evaluation of its role as a single screening modality in the Canadian National Breast Screening Study. *Cancer* 1989; 63: 160-166
15. Miller AB: The Canadian National Breast Screening Study. In Day NE, Miller AB (eds): *Screening for Breast Cancer*, Huber, Toronto, 1988: 51-58
16. Miller AB, Baines CJ, To T et al: The Canadian National Breast Screening Study. In Miller AB, Chamberlain J, Day NE et al (eds): *Cancer Screening*, Cambridge U Pr, Cambridge, Engl, 1991: 45-55
17. Yaffe M, Mawdsley GE, Nishikawa RM: Quality assurance in a national breast screening study. *Proc Soc Photo-Optical Instrum Eng* 1983; 419: 23-30
18. Baines CJ, McFarlane DV, Wall C: Audit procedures in the National Breast Screening Study: mammography interpretation. *J Can Assoc Radiol* 1986; 39: 256-260
19. Baines CJ, McFarlane DV, Miller AB: The role of the reference radiologist: estimates of inter-observer agreement and potential delay in cancer detection in the National Breast Screening Study. *Invest Radiol* 1990; 25: 971-976
20. Miller AB, Baines CJ, Turnbull C: The role of the nurse-examiner in the National Breast Screening Study. *Can J Public Health* 1991; 82: 162-167
21. Baines CJ, Miller AB, Kopans DB et al: Canadian National Breast Screening Study: assessment of technical quality by external review. *AJR* 1990; 155: 743-747
22. Baines CJ, Christen A, Simard A et al: The NBSS: pre-recruitment sources of awareness. *Can J Public Health* 1989; 80: 221-225
23. Morrison AS, Brisson J, Khalid N: Breast cancer incidence and mortality in the Breast Cancer Detection Demonstration Project. *J Natl Cancer Inst* 1988; 80: 1540-1547
24. Cutler SJ, Ederer F: Maximum utilization of the life table method in analyzing survival. *J Chronic Dis* 1958; 8: 699-712
25. Peto R, Pike MC, Armitage P et al: Design and analysis of randomised clinical trials requiring prolonged observation of each patient: II. Analysis and examples. *Br J Cancer* 1977; 35: 1-39
26. *1981 Census of Canada* (public use sample tape 8-1200-609), Statistics Canada, Ottawa
27. Schechter MT, Miller AB, Baines CJ et al: Selection of women at high risk of breast cancer for initial screening. *J Chronic Dis* 1986; 39: 253-260
28. Holowaty PH, Miller AB, Baines CJ et al: Risk factors for breast cancer occurring after an initial screen in the Canadian National Breast Screening Study. *Cancer Epidemiol Biomarkers Prevent* (in press)

29. Simard A, Paquette L, Baillargeon J et al: Perception of cancer detection and early treatment in a population participating in the National Breast Screening Study in Canada. *Can J Public Health* 1989; 80: 226-227
30. Day NE, Baines CJ, Chamberlain J et al: UICC project on screening for cancer: report of the Workshop on Screening for Breast Cancer. *Int J Cancer* 1986; 38: 303-308
31. Howe GR, Friedenreich CM, Jain M et al: A cohort study of fat intake and risk of breast cancer. *J Natl Cancer Inst* 1991; 83: 336-340
32. Brisson J, Verreault R, Morrison AS et al: Diet, mammographic features of breast tissue, and breast cancer risk. *Am J Epidemiol* 1989; 130: 14-24
33. Shapiro S, Strax P, Venet L et al: The search for risk factors in breast cancer. *Am J Public Health* 1968; 58: 820-835
34. Baker LH: Breast Cancer Detection Demonstration Project: five-year summary report. *CA Cancer J Clin* 1982; 32: 194-225
35. Miller AB, Baines CJ, Sickles EA: Canadian National Breast Screening Study [C]. *AJR* 1990; 155: 1133-1134
36. Tabár L, Akerlund E, Gad A: Five-year experience with single-view mammography randomized controlled trial in Sweden. *Recent Results Cancer Res* 1984; 90: 105-113
37. United Kingdom Trial of Early Detection of Breast Cancer Group: Specificity of screening in United Kingdom Trial of Early Detection of Breast Cancer. *BMJ* 1992; 304: 346-349
38. Tabár L, Fagerberg G, Day NE et al: What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. *Br J Cancer* 1987; 55: 547-551
39. Verbeek ALM, Hendriks JHCL, Holland R et al: Mammographic screening and breast cancer mortality: age-specific effects in Nijmegen Project, 1975-82 [C]. *Lancet* 1985; 1: 865-866
40. Andersson I, Aspegren K, Janzon L et al: Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *BMJ* 1988; 297: 943-948
41. Tabár L, Fagerberg CJG, Gad A et al: Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985; 1: 829-832
42. Frisell J, Eklund G, Hellström L et al: Randomized study of mammographic screening — preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat* 1991; 18: 49-56
43. Palli D, del Turco R, Buiatti E et al: A case-control study of the efficacy of a non-randomized breast cancer screening program in Florence (Italy). *Int J Cancer* 1986; 38: 501-504
44. Roberts MM, Alexander FE, Anderson TJ et al: Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet* 1990; 335: 241-246
45. Kerr M: *A Case-Control Study of Treatment Adequacy and Mortality from Breast Cancer for Women of Age 40-49 Years at Entry into the National Breast Screening Study* (master's thesis), U of Toronto, Toronto, 1991

Conferences

continued from page 1454

Mar. 14-18, 1993: 4th International Conference on the Reduction of Drug Related Harm (sponsored by the Dutch Ministry of Welfare, Health and Cultural Affairs, the Municipality of Rotterdam and the *International Journal on Drug Policy*)
Rotterdam, the Netherlands
Conference Secretariat, Essenlaan 16, PO Box 4193, 3006 AD, Rotterdam, the Netherlands; tel 011-31-20-0-10-452-51-66, fax 011-31-20-0-10-452-07-71

Mar. 18-23, 1993: Association for Applied Psychophysiology and Biofeedback 24th Annual Meeting
Los Angeles
Joette Cross, director of meetings, Association for Applied Psychophysiology and Biofeedback, Ste. 304, 10200 W 44th Ave., Wheat Ridge, CO 80033; tel (303) 422-8436, fax (303) 422-8894

Mar. 19-21, 1993: Canadian-Trinidad and Tobago Medical Convention (sponsored by the Trinidad and Tobago Medical Association)
Trinidad and Tobago
Medical Staff Office, Queensway General Hospital, 150 Sherway Dr., Etobicoke, ON M9C 1A5; tel (416) 253-2938, fax (416) 253-0111

Apr. 4-8, 1993: 13th World Congress on Occupational Safety and Health
New Delhi, India
Official languages: English, French, Spanish, German and Japanese
Congress Secretariat, 13th World Congress on Occupational Safety and Health, National Safety Council, PO Box 26754, CLI Building — Sion, Bombay 400 022, India; tel 011-91-22-4073285, fax 011-91-22-4075937

Apr. 5-9, 1993: 4th International Meeting on Trace Elements in Medicine and Biology — Trace Elements and Free Radicals in Oxidative Diseases (organized by the Society for Free Radical Research and the Société francophone d'étude et de recherche sur les éléments trace essentiels)
Chamonix, France
Official language: English. Simultaneous translation languages: French-English.
Prof. Alain Favier or Mme. Arlette Alcaraz, Laboratoire de Biochimie C, Hôpital A. Michallon, BP 217X, 38043 Grenoble Cédex 09, France; tel 011-33-76-76-54-07, fax 011-33-76-42-66-44

Apr. 18, 1993: 5th Annual Symposium on Treatment of Headaches and Facial Pain
New York
Dr. Alexander Mauskop, Director, New York Headache Center, 301 E 66th St., New York, NY 10021; tel (212) 794-3550

continued on page 1498